Journal of Institutional and Theoretical Economics

# JITE

Vol. 171, No. 3 September 2015



ARTIBUS INGENUIS J·C·B·M· 1·8·0·1

Mohr Siebeck

*JITE* is published 4 times a year by
**Mohr Siebeck GmbH & Co. KG, P.O. Box 2040, 72010 Tübingen, Germany**

*JITE* is printed on acid-free and permanent/book paper ♾
meeting all library standards including ANSI/NISO Z39.48-1988.

**Papers for Publication**
must be in English. They should be submitted online at https://editorialexpress.com/jite.
Editorial office: Prof. Dr. Gerd Muehlheusser, Department of Economics, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany,
E-mail: jite@mohr.de.

Please follow the Instructions for Authors on the inside back cover.

There is no submission fee. Submission of a paper will be held to imply that it contains original unpublished work and is not being submitted for publication elsewhere.

The Editor does not accept responsibility for damage or loss of paper submitted.
Articles are refereed by specialists.

**Books for Review**
should be addressed to the Publisher.

**Transfer of Rights**
When an article is accepted for publication, the exclusive copyright is granted to Mohr Siebeck for publication in a print and an electronic version. Further information on this and the rights retained by the author can be found at www.mohr.de/jite. No one may reproduce or distribute the entire journal or parts of it in a print or an electronic version without the publisher's permission. Please contact rights@mohr.de.

**Subscriptions**
*JITE* is published in quarterly issues of approximately 750 pages per year. Back volumes are available on request.
New orders, renewals, claims, and all other correspondence relating to subscriptions should be addressed to the Journals Marketing Manager at the Publisher's address above
(Fax +49 7071 51104 or by e-mail to journals@mohr.de). Please quote the name of the journal.

**Full Text Online**
Free access to the full text online is included in a subscription. We ask institutions with more than 20.000 users to obtain a price quote directly from the publisher. Contact: sandra.witt@mohr.de.
In order to set up Online access for institutions / libraries, please go to:
http://www.ingentaconnect.com/register/institutional.
In order to set up Online access for private persons, please go to:
http://www.ingentaconnect.com/register/personal.



**ISSN 0932-4569**

# On Repeated Games with Endogenous Matching Decision

by

Heiner Schumacher*

Received August 5, 2012; in revised form February 28, 2015;
accepted March 26, 2015

We study infinitely repeated games that are played by many groups simultaneously and where players have the option to maintain or quit relationships. For two-player stage games any individually rational payoff vector in the relative interior of $V^*$ can be sustained as equilibrium payoff if the discount factor $\delta$ is sufficiently large. Such a statement is not possible for stage games with more than two players. We translate the refinement of weak renegotiation-proofness to our framework and characterize the set of payoffs that can be sustained through strategies that are "bilaterally rational" in the sense of Ghosh and Ray (1996). (JEL: C70, C72)

## 1 Introduction

Most of the literature on infinitely repeated games or relational contracts considers settings with exogenously given matching protocols: Either a player is forced to play against a fixed set of opponents all the time, or she plays against different opponents in each period (e.g., Kandori, 1992, or Ellison, 1994). However, in almost all social and economic interactions individuals are free to quit relationships and establish new ones. In this paper, we therefore study the equilibrium set in infinitely repeated games with the option to maintain or to quit relationships.

The setup is as follows. An $N$-player normal-form game is played simultaneously in $M$ groups. After observing the opponents' action choice, each player can choose whether to maintain the relationship with her current group or not. If at least one player of a group quits the relationship, all players of this group return to a pool of unmatched players, from which new groups are formed randomly at the beginning

* Aarhus University, Denmark. I thank Biung-Ghi Ju, two anonymous referees, Francesco Squintani for supervision, and V. Bhaskar, Markus Kinateder, Ludwig Rener, and Ernst-Ludwig von Thadden for helpful suggestions. Financial support by the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged.

*Journal of Institutional and Theoretical Economics* 171, 544–564 – ISSN 0932-4569
DOI: 10.1628/093245615X14302136524716 – 

of the next period. With probability $1 - \delta$ a player "dies" and will be replaced in the next period by a new player who joins the pool of unmatched players. With probability $\delta$ she survives until the next period. If all players of a group maintain the relationship and survive the period, they play the stage game against each other again. A player only observes the identity of current and former opponents and the actions these players have chosen in the periods in which she played the stage game with them. There are no information flows between groups.

Similar settings have been analyzed in a number of papers; see Datta (1993), Kranton (1996), Ghosh and Ray (1996), Carmichael and MacLeod (1997), Fujiwara-Greve and Okuno-Fujiwara (2009), Rob and Yang (2010), and Schumacher (2013).[1] The focus of these papers is cooperation in the two-player prisoners' dilemma. To sustain cooperation in equilibrium players can *start small*: At the beginning of a new relationship, players defect and start to cooperate in later periods.[2] Whenever a player deviates from this path of play, her opponent quits the relationship. Thus, any gain from deviation is wiped out by the subsequent phase of low payoffs in the next relationship.[3]

In contrast to the previous literature, we allow for any finite stage game. Our object of interest is the set of payoff vectors that can be supported as expected normalized equilibrium payoffs for all players, regardless of the number $M$ of groups. Recall that any individually rational payoff vector $v \in V^*$ can be sustained in a subgame-perfect equilibrium of the canonical infinitely repeated game if the discount factor $\delta$ is sufficiently large (Fudenberg and Maskin, 1986). This folk theorem does not hold in our framework. However, by using starting-small strategies, we show that any payoff vector in the relative interior of those payoffs, if it strictly dominates a convex combination of Nash payoffs of the stage game, can be supported as expected normalized equilibrium payoff if $\delta$ is sufficiently large. Our first main result is an almost complete characterization of the set of equilibrium payoffs for two-player stage games. If $N = 2$, any individually rational payoff vector $v$ in the relative interior of $V^*$ can be sustained in an equilibrium, provided that $\delta$ is sufficiently large. Hence, for two-player stage games and large discount factors $\delta$, the set of equilibrium payoffs in the infinitely repeated game with the option to maintain or to quit relationships is almost identical to the one in the canonical repeated game. In contrast, for stage games with more than two player roles, the option to quit relationships may greatly reduce the set of equilibrium payoffs. We

[1] The option to maintain or to quit relationships has also been considered in other settings; see Matsushima (1990) or Casas-Arce (2010). Jackson and Watts (2008, 2010) analyze the existence of equilibria in finitely repeated games where players choose both actions and opponents.

[2] Starting-small strategies have also been analyzed experimentally; see Andreoni and Samuelson (2006).

[3] Fujiwara-Greve and Okuno-Fujiwara (2009) show that there also can be an equilibrium where a fraction of players start to cooperate immediately (and quit the relationship whenever the opponent defects) and all other players choose a starting-small strategy with one period of defection.

show by example that individually rational payoffs close to the minmax payoff profile are not always equilibrium payoffs.

We then turn to the renegotiation-proofness of starting-small strategies. Ghosh and Ray (1996) argue that starting-small strategies may not be robust against joint deviations, since players have an incentive to drop the phase of low payoffs at the beginning of a relationship and start cooperation immediately. In several papers, this problem is solved by assuming incomplete information about players' types. For example, a player may be patient or myopic; see Ghosh and Ray (1996) or Rob and Yang (2010). We show that even in games with complete information and homogeneous time preferences, starting-small strategies can be robust to joint deviations. Following Farrell and Maskin (1989), we construct equilibrium strategies in which, at the beginning of a new relationship (and after any unilateral deviation), a profile is played that punishes one player and rewards her opponent. Our second main result is a characterization of payoff profiles that can be sustained in a "weakly group-specific renegotiation-proof equilibrium" (below we define this term formally).

The rest of the paper is organized as follows. Section 2 introduces the basic framework. Section 3 provides a characterization of the set of equilibrium payoffs that holds for any number $M$ of groups. In section 4, we adapt Farrell and Maskin's (1989) refinement of weak renegotiation-proofness to our framework. Section 5 concludes. All proofs can be found in the appendix.

## 2 *Framework of the Model*

Time is discrete and denoted by $t \in \{0, 1, ...\}$. In each period, an $N$-player normal-form game $G$ is played in $M \in \mathbb{N}$ groups. In each group, there is exactly one player of each player role $i \in \{1, ..., N\}$. There is no discounting. However, at the end of period $t$, each player dies with probability $1 - \delta$ and is replaced by a new player who takes on the same role. With probability $\delta$ she survives the period.[4] A player's identity is denoted by $i \cdot m \cdot g$, where $g$ represents her generation. If player $i \cdot m \cdot g$ dies, she is replaced by player $i \cdot m \cdot (g + 1)$. In period 0, we have $g = 1$ for all players.

At the beginning of a period, each player is either matched in a group or unmatched. All unmatched players are matched randomly into groups.[5] Thus, all players who are alive in a given period play $G$ in some group. Denote the group of player $i \cdot m \cdot g$ in period $t$ by the set $r^t_{i \cdot m \cdot g}$, which includes $i \cdot m \cdot g$ and the identities

[4] Thus, $\delta$ can be interpreted as the usual discount factor. Including both discounting and random death would not affect our results.

[5] By construction, there are the same number of players of each role in the pool of unmatched players. All possible ways of pairing up these players have the same probability. The assumption of random matching is a shortcut to the formation of groups through partner choice. Our results would not change if players could avoid being matched to previous opponents when there are others in the pool of unmatched players.

of her opponents in this period. We say that player $i \cdot m \cdot g$ starts in a new group in period $t$ if $r^t_{i \cdot m \cdot g} \neq r^{t-1}_{i \cdot m \cdot g}$.

In each period, players in role $i$ choose an action from the pure finite action set $A_i$. Denote the action chosen by player $i \cdot m \cdot g$ in period $t$ by $a^t_{i \cdot m \cdot g}$. The set of mixed actions for players in role $i$ is $\Delta(A_i)$; the set of pure action profiles is $A \equiv \prod_{i=1}^N A_i$ with typical element $a$; and the set of mixed action profiles is $\Delta(A) \equiv \prod_{i=1}^N \Delta(A_i)$ with typical element $\alpha$. Mixed actions are not observable. The action profile realized in the group of $i \cdot m \cdot g$ in period $t$ is given by $a^{t,i \cdot m \cdot g}$.[6] Stage-game payoffs are given by a function $u : A \to \mathbb{R}^N$. Let $u(\alpha)$ be the vector of expected payoffs from a mixed action profile $\alpha$.

After the action profile is realized, each player has the option to maintain ($MT$) or to quit ($Q$) the relationship with her current group. If at least one player of a given group chooses $Q$ or if at least one player of this group dies, all members of this group join the pool of unmatched players (together with the new players) at the beginning of the next period. Otherwise, all players of this group remain matched together for one more period. The sequence of events in each period is as follows: (i) Unmatched players are matched randomly into groups; (ii) all players choose their actions, and payoffs are realized; (iii) each player chooses between $MT$ and $Q$; (iv) players die and are replaced by new players with probability $1 - \delta$, groups in which all players have chosen $MT$ and survive the period stay together, and all other players enter the pool of unmatched players together with the new ones.

Players recall the history of play in their own groups. In each period $t$, they also observe the realization $w^t$ of a public correlation device before matching occurs, i.e., before stage (i). Assume that $w^t$ is uniformly distributed on $[0, 1]$. Let $t_{i \cdot m \cdot g}$ be the period in which player $i \cdot m \cdot g$ plays $G$ for the first time (her *period of birth*). The history of play for player $i \cdot m \cdot g$ in a period $t > t_{i \cdot m \cdot g}$ is then given by

$$h^t_{i \cdot m \cdot g} = (a^{\tau,i \cdot m \cdot g}, w^\tau)_{\tau \in \{t_{i \cdot m \cdot g}, \ldots, t-1\}} .$$

Let H be the set that includes all possible finite histories of play, including the empty history. The history of groups for this player is given by

$$R^t_{i \cdot m \cdot g} = \left(r^\tau_{i \cdot m \cdot g}\right)_{\tau \in \{t_{i \cdot m \cdot g}, \ldots, t\}} .$$

Let R be the set that includes all possible finite histories of groups, including the empty history. The strategy $\sigma_i$ of players in role $i$ consists of two parts, action choice and matching decision. The action choice is a function that maps the period of birth, the identity (i.e., the numbers $m$ and $g$), the history of play, the history of groups, and the current realization of the public signal into the set of mixed actions,

$$\sigma_i^{[1]} : \mathbb{N}^3 \times \mathrm{H} \times \mathrm{R} \times [0, 1] \to \Delta(A_i) .$$

The matching decision is a function that maps the period of birth, the identity, the history of play, the history of groups, the current realization of the public signal, and the realized action profile into the set of mixed matching decisions,

$$\sigma_i^{[2]} : \mathbb{N}^3 \times \mathrm{H} \times \mathrm{R} \times [0, 1] \times A \to \Delta(\{Q, MT\}) .$$

[6] We will sometimes suppress the notation for individual players.

Note that each player in role $i$ may act differently from all the others. For a given strategy profile $\sigma = \{\sigma_i\}_{i \in \{1,\dots,N\}}$ one can calculate the expected normalized payoff of player $i \cdot m \cdot g$ at her birth, which is

$$(1-\delta)\mathrm{E}_\sigma \left[ \sum_{t=t_{i \cdot m \cdot g}}^{\infty} \delta^{t - t_{i \cdot m \cdot g}} u_i(a^{t, i \cdot m \cdot g}) \mid t_{i \cdot m \cdot g}, m, g \right].$$

As a player only observes the action profiles realized in her own groups, we have a game of imperfect information. Our solution concept is sequential equilibrium. A sequential equilibrium requires a strategy profile and a system of beliefs that are sequentially rational and consistent, respectively. We will restrict attention to a class of strategy profiles in which every player conditions her action choices only on her observations. Each of our strategy profiles is sequentially rational for any belief a player may have about the unobserved actions chosen by players in other groups. Hence, beliefs are not modeled explicitly, and a sequential equilibrium is said to exist when the condition of sequential rationality is fulfilled. The set of payoffs generated by pure action profiles is

$$V \equiv \left\{ v \in \mathbb{R}^N : \exists\, a \in A \;\text{ s.t. } v = u(a) \right\},$$

while the set of feasible payoffs $V^\dagger$ is given by the convex hull of $V$, i.e., $V^\dagger \equiv \mathrm{co}(V)$. The min–max payoff for a player in role $i$ is given by

$$\hat{v}_i = \min_{\alpha_{-i} \in \prod_{j \neq i} \Delta(A_j)} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i});$$

the corresponding min–max profile is $\hat{\alpha}^i$. The set of individually rational payoffs is

$$V^* \equiv \left\{ v \in V^\dagger : v_i > \hat{v}_i, i = 1, ..., N \right\}.$$

Let $\Sigma^{NE} \subset \Delta(A)$ be the set of all Nash equilibria of $G$. By Nash's (1951) existence theorem, this set has at least one element. The set of payoffs generated by Nash equilibria is

$$V_{NE} \equiv \left\{ v \in \mathbb{R}^N : \exists\, \alpha \in \Sigma^{NE} \;\text{ s.t. } v = u(\alpha) \right\},$$

while the set of feasible Nash payoffs $V_{NE}^\dagger$ is the convex hull of $V_{NE}$, $V_{NE}^\dagger \equiv \mathrm{co}(V_{NE})$. The set of feasible payoffs that strictly dominate a convex combination of Nash payoffs is

$$V_{NE}^* \equiv \left\{ v \in V^\dagger : \exists\, \bar{v} \in V_{NE}^\dagger \;\text{ s.t. } v_i > \bar{v}_i, i = 1, ..., N \right\}.$$

Finally, define role $i$'s cheating payoff from the action profile $a$ by

$$c_i(a) = \max_{\hat{a}_i \in A_i} u_i(\hat{a}_i, a_{-i}), \quad v_i^{\max} = \max_{a \in A} u_i(a), \quad \text{and} \quad v_i^{\min} = \min_{a \in A} u_i(a).$$

## *3 A Folk Theorem for Any Number of Groups*

In a finite population, there is always a positive probability of being paired up with the same opponents after choosing $Q$ in the preceding period. Suppose that all players choose $MT$ in each period. An increase in $\delta$ then has two effects: first, it increases the value of future payoffs, and second, it increases the probability of being paired up with the same opponents again after choosing $Q$ (because it is less likely that a player from another group dies). As players can condition their actions on the identity of their opponents, there is scope for punishment within a group. It is then simple to establish the folk theorem of Fudenberg and Maskin (1986) for our framework: For given $M$, if the dimension of $V^{\dagger}$ equals $N$, then for any $v \in V^*$ there is a value $\bar{\delta} < 1$ such that there is a sequential equilibrium in which $v_i$ is the expected normalized payoff for each player in role $i$ whenever $\delta \geq \bar{\delta}$.[7]

The finite size of the population may also allow for cooperation through *contagious strategies* as in Ellison (1994). Suppose that the stage game $G$ is the prisoners' dilemma, and that all players always choose $Q$. Then in each period all players are rematched, and no player can maintain a relationship by unilaterally choosing $MT$ (hence, choosing $Q$ is optimal). A contagious strategy works as follows: if a player defects in period $t$, then she continues defecting, her opponent starts defecting in period $t+1$, the opponent's opponent starts defecting in period $t+2$, and so forth. Thus, a single deviation triggers a breakdown of cooperation, which punishes the initial deviator. For given $M$, if $\delta$ is sufficiently large, the threat of breakdown may induce players to cooperate in each period.

However, we are interested in the set of equilibrium payoffs when the population is so large that it is unlikely to be matched to the same opponents after choosing $Q$, and contagious strategies no longer work because contagion would be too slow to deter defection. We can show formally that for given $\delta$, if $M$ is sufficiently large, no equilibrium exists in which a profile that is not Nash is played in all groups and periods.

LEMMA 1 *For given $\delta < 1$, there is $\bar{M} \in \mathbb{N}$ such that there is no equilibrium in which all groups play a profile $\alpha \notin \Sigma^{NE}$ in all periods if $M > \bar{M}$.*

In the prisoners' dilemma, the payoff from mutual cooperation is not an equilibrium payoff for given $\delta$ if $M$ is sufficiently large. Thus, for many stage games, $V^*$ is no longer the set of equilibrium payoffs for large $\delta$ and arbitrary $M$.

Nevertheless, there is at least one possibility to support the repeated play of profiles that are not Nash equilibria of the stage game, regardless of $M$. Suppose that a stage game has two action profiles $\alpha^1 \in \Sigma^{NE}$ and $a^2 \in A \backslash \Sigma^{NE}$ with $u_i(\alpha^1) < u_i(a^2)$ for all $i$. Let $T \in \mathbb{N}$, and consider following strategy for player $i \cdot m \cdot g$: Always choose $MT$; in period $t$, choose $a_i^2$ if $r_{i \cdot m \cdot g}^{t-1} = r_{i \cdot m \cdot g}^{t}$ and $a^{t-1, i \cdot m \cdot g} = a^2$ or if $r_{i \cdot m \cdot g}^{\tau} = r_{i \cdot m \cdot g}^{t}$ and $a_{i \cdot m \cdot g}^{\tau} = \alpha_i^1$ in all periods $\tau \in \{t-T, ..., t-1\}$; otherwise, choose $a_i^t = \alpha_i^1$.

This is a version of the starting-small strategy mentioned in the introduction: when the members of a new group are matched together, a Nash profile $\alpha^1$ is played for $T$ periods. Call this the *punishment phase*. Afterwards, the group plays a profile $a^2$, which is more favorable for all players. Call this the *normal phase*. The punishment phase is also triggered when some player in role $i$ deviates from $a_i^2$. If all players in

[7] The assumption of full dimensionality can be replaced by the weaker assumption of *nonequivalent utilities* (Abreu, Dutta, and Smith, 1994).

all groups play according to this strategy and $T$ and $\delta$ are sufficiently large, then any gain from deviation is wiped out by the subsequent punishment phase, regardless of whether the deviator is matched to the same group or not.

Formally, we say that an equilibrium $\sigma$ consists of starting-small strategies if it has the following properties. Players always choose $MT$. There are two action profiles (or public randomizations) $\alpha^1$ and $\alpha^2$ as well as some $T \in \mathbb{N}$ such that on the equilibrium path each player in role $i$ who starts in a new group chooses $\alpha_i^1$ in the first $T$ periods of this group, and $\alpha_i^2$ in each period thereafter until the group is broken up.

By using starting-small strategies, we can establish a partial characterization of the equilibrium payoffs that holds for any number $M$ of groups.

LEMMA 2 *For any $v$ in the relative interior*[8] *of $V_{NE}^*$, there is a value $\bar{\delta} < 1$ such that if $\delta > \bar{\delta}$, then there is a sequential equilibrium in which the expected normalized payoff for each player in role i equals $v_i$.*

Starting-small strategies may support an equilibrium even if the punishment-phase profile is not a Nash equilibrium of the stage game and the period payoff in the normal phase does not dominate a convex combination of Nash payoffs. However, if the punishment-phase profile is not a Nash equilibrium, we additionally have to rule out that some players can profitably deviate in the punishment phase. Consequently, not every action profile with an individually rational payoff vector can be used as normal-phase profile in general. This makes a complete characterization of all equilibrium payoff vectors for arbitrary $M$ difficult and reduces the set of equilibrium payoffs. An example illustrates this point.

*Example 1.* Consider the three-player stage game in Figure 1, in which a player in role 1 chooses rows, a player in role 2 chooses columns, and a player in role 3 chooses matrices. The min–max payoffs are $\hat{v}_1 = \hat{v}_2 = \hat{v}_3 = 0$, and thus, a payoff of 1 is individually rational for each player. We show that no equilibrium $\sigma$ achieves this payoff profile for arbitrary $M$.

*Figure 1*
Stage Game from Example 1

| $X$ | $L$ | $C$ | $R$ |
|---|---|---|---|
| $T$ | 10,10,10 | 10,1,1 | 0,0,0 |
| $P$ | 1,1,1 | 0,10,0 | 0,10,0 |
| $B$ | 0,0,0 | 0,0,0 | 0,0,0 |

| $Y$ | $L$ | $C$ | $R$ |
|---|---|---|---|
| $T$ | 10,10,1 | 10,1,1 | 10,10,10 |
| $C$ | 1,1,10 | 1,10,0 | 10,10,10 |
| $B$ | 1,1,10 | 1,1,10 | 1,1,10 |

[8] Recall that for the affine hull $\operatorname{aff}(V_{NE}^*) = \{\sum_{j=1}^n \lambda_j u(a^{[j]}) : n \in \mathbb{N}, a^{[j]} \in A, \sum_{j=1}^n \lambda_j = 1\}$ the relative interior of $V_{NE}^*$ is those points $v \in \operatorname{aff}(V_{NE}^*)$ for which there exists an $\varepsilon > 0$ such that if $v' \in \operatorname{aff}(V_{NE}^*)$ and $d(v, v') \leq \varepsilon$, then $v' \in V_{NE}^*$, where $d$ is the Euclidean distance.

As a preliminary step, we define a *virgin group* for $i \cdot m \cdot g$ (with $g \geq 2$) as a group $r$ in which (i) all roles are taken by players in generation 2 or later, (ii) no two players interacted with each other in previous periods, and (iii) no sequence of players $x_1, x_2, ..., x_n$ with $x_n \in r$ exists such that $i \cdot m \cdot g$ has interacted with $x_1$, $x_1$ has interacted with $x_2$, ..., and $x_{n-1}$ has interacted with $x_n$. Clearly, if $M$ and $t$ become large, the probability that $i \cdot m \cdot g$ is matched to a virgin group for her after choosing $Q$ converges to 1.

Now assume by contradiction that an equilibrium $\sigma$ exists in which the normalized expected payoff for all players equals 1. In a virgin group (for some player), the expected normalized payoff of each player must be 1 (because with positive probability she was just born and then by definition must earn 1 in expectation). The probability that an action is played in this group that yields some player a payoff of 10, say $i \cdot m \cdot g$, cannot exceed $1/10$. Otherwise, if $M$ and $t$ are sufficiently large, player $i \cdot m \cdot g$ could choose $Q$ (to get matched to another virgin group for her) and thereby increase her expected normalized payoff. Hence, with a probability of at least $7/10$ one of the remaining actions must be played, i.e., $\{T, R, X\}$, $\{P, L, X\}$, $\{B, L, X\}$, $\{B, C, X\}$, or $\{B, R, X\}$. However, when such an action is played, the player in role 3 can increase her period payoff from 0 or 1 to 10 by choosing $Y$. If $M$ and $t$ are sufficiently large, this player can secure herself an expected normalized payoff above 1 by choosing $Y$ and $Q$, a contradiction.

For two-player games individual rationality becomes meaningful again. By playing the mutual min–max profile $\{\hat{\alpha}_1^2, \hat{\alpha}_2^1\}$, both players earn a period payoff that is strictly below any individually rational payoff. Also, by choosing a best response to $\hat{\alpha}_{-i}^i$, a player in role $i$ earns $\hat{v}^i$, which again is strictly below any individually rational payoff. Using this fact, we get an almost complete characterization of all equilibrium payoffs.

PROPOSITION 1 *Assume that $N = 2$. For any $v$ in the relative interior of $V^*$, there is a value $\bar{\delta} < 1$ such that if $\delta > \bar{\delta}$, then there is a sequential equilibrium in which the expected normalized payoff for each player in role $i$ equals $v_i$.*

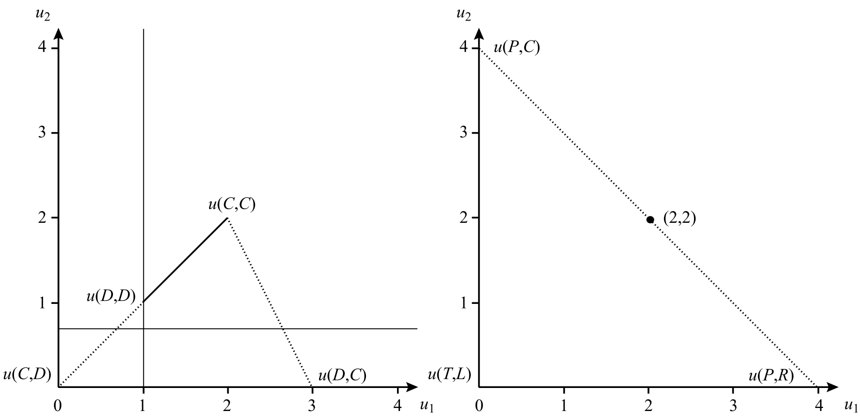
*Example 2.* Even for two-player games we cannot hope for a simple *and* complete characterization of equilibrium payoffs for arbitrary $M$. Consider the stage game $G_1$ in Figure 2. A player in role 1 chooses rows, and a player in role 2 chooses columns.

*Figure 2*
Stage Game from Example 2 ($G_1$)

| $G_1$ | $D$ | $C$ |
|---|---|---|
| $D$ | 1,1 | 3,0 |
| $C$ | 0,0 | 2,2 |

The unique Nash equilibrium of the stage game is $\{D, D\}$, and the min–max payoffs are $\hat{v}_1 = 1$ and $\hat{v}_2 = 2/3$. The convex combinations of $u(\{D, D\})$ and $u(\{C, C\})$

*Figure 3*
The Set of Feasible and Individually Rational Payoffs in $G_1$ (left) and $G_2$ (right)




are located at the boundary of $V^*$ (see Figure 3). By constructing a starting-small strategy with $\{D, D\}$ as punishment-phase profile and a public randomization (using $\{D, D\}$ and $\{C, C\}$) as normal-phase profile, we can establish any of these convex combinations – except $u(\{C, C\})$ – as equilibrium payoff for all players, provided that $\delta$ is sufficiently large. Hence, individually rational payoff profiles $v \notin V^{\dagger}_{NE}$ located at the boundary of $V^*$ can be equilibrium payoffs for all players and arbitrary $M$.

Next, consider the stage game $G_2$ in Figure 4, where a player in role 1 chooses rows and a player in role 2 chooses columns. The min–max payoffs are $\hat{v}_1 = \hat{v}_2 = 0$. The convex combinations of $u(\{P, C\})$ and $u(\{P, R\})$ are located at the boundary of $V^*$ (see Figure 3). One of these convex combinations (a payoff of 2 for each player) is the payoff of a Nash equilibrium of the stage game where the player in role 1 (role 2) chooses $P$ and $B$ ($C$ and $R$) with equal probability. We show that none of the other convex combinations is an equilibrium payoff profile for arbitrary large $M$.

*Figure 4*
Stage Game from Example 2 ($G_2$)

| $G_2$ | $L$ | $C$ | $R$ |
|---|---|---|---|
| $T$ | 0,0 | 0,0 | 0,0 |
| $P$ | 0,0 | 0,4 | 4,0 |
| $B$ | 0,0 | 4,0 | 0,4 |

Assume by contradiction that an equilibrium $\sigma$ exists in which the expected normalized payoffs of players in roles 1 and 2 are $v_1 \neq 2$ and $4 - v_1$, respectively.

Without loss of generality we assume $v_1 < 2$. Since the expected normalized payoff profile is a linear combination of the stage-game payoffs, an action profile with period payoffs (4, 0) or (0, 4) must be played in all periods and groups. Thus, on the equilibrium path, players in role 1 (role 2) always choose action $P$ or $B$ ($C$ or $R$).

Consider the following alternative strategy for a player in role 1 and $g \geq 2$: In each period, she plays a best response to the profile that is supposed to be played in her current group and quits the relationship. If $M$ and $t$ are sufficiently large, she mostly plays in virgin groups for her, in which, by definition, the opponent in role 2 chooses $C$ or $R$. In these groups, she earns a period payoff of at least 2 in expectation. Hence, for $M \to \infty$ and $t \to \infty$ her expected normalized payoff from the alternative strategy converges to $2 > v_1$, a contradiction. We conclude that feasible and individually rational payoff profiles $v$ located at the boundary of $V^*$ may or may not be equilibrium payoffs for all players and arbitrary $M$, depending on the structure of the stage game.

## 4 *Renegotiation-Proofness*

If we consider the punishment phase of a starting-small strategy, it is evident that players who have just been matched together in a new group may want to drop the punishment and start with the normal phase immediately. Given that all other players stick to the starting-small strategy, no player of this group could increase her payoff by deviating in the normal phase. Consider, for example, the prisoners' dilemma $G_3$ in Figure 5, in which a player in role 1 chooses rows, and a player in role 2 chooses columns.

*Figure 5*
Prisoners' Dilemma $G_3$

| $G_3$ | $D$ | $C$ |
|---|---|---|
| $D$ | 1,1 | 5,0 |
| $C$ | 0,5 | 4,4 |

A starting-small strategy with mutual defection $\{D, D\}$ as punishment-phase profile and mutual cooperation $\{C, C\}$ as normal-phase profile sustains an equilibrium, provided that $\delta$ is sufficiently large. If a newly matched group $r$ starts with the normal phase immediately and complies with the starting-small strategy in all future periods, these players increase their expected discounted payoff by 3. Since all other groups play according to the starting-small strategy (now and in the future), the players of $r$ cannot deviate profitably from this agreement. Hence, it is credible. However, if all groups always drop the punishment phase, we are no longer in equilibrium.

Ghosh and Ray (1996) call this problem of the starting-small strategy a lack of *bilateral rationality*. Under complete information, equilibria in starting-small strategies may not be robust to Pareto-improving and incentive-compatible joint deviations.

In this section, we examine the circumstances under which a starting-small strategy exists that sustains the repeated play of profiles $\alpha \notin \Sigma^{NE}$ in equilibrium and is robust to this criticism. To capture the concept of bilateral rationality formally, we adopt Farrell and Maskin's (1989) refinement of weak renegotiation-proofness for two-player games to our setting.[9] This refinement seems to be related to bilateral rationality: It rules out strategy profiles $\sigma$ (as defined in the canonical repeated game) in which some continuation equilibrium is Pareto-dominated by another continuation equilibrium of $\sigma$. To illustrate this, consider again the prisoners' dilemma $G_3$. Suppose that two players in the canonical repeated game play the grim-trigger strategy "Choose $C$ unless at least one player has chosen $D$ in the past; in this case, choose $D$." For sufficiently large $\delta$, this strategy sustains cooperation in equilibrium. However, the phase of mutual defection punishes both players. Thus, they may agree to start the game anew with mutual cooperation. The grim-trigger strategy is therefore not robust to renegotiations. Formally, weak renegotiation-proofness is defined as follows:

DEFINITION 1 (WEAK RENEGOTIATION-PROOFNESS; FARRELL AND MASKIN, 1989) *A subgame-perfect equilibrium $\sigma$ (as defined for the canonical repeated game) is weakly renegotiation-proof if there do not exist continuation equilibria $\sigma^1$, $\sigma^2$ of $\sigma$ such that $\sigma^1$ strictly Pareto-dominates $\sigma^2$.*

Farrell and Maskin (1989) construct weakly renegotiation-proof equilibria by using profiles that punish the deviator and benefit her opponent. Consider again the prisoners' dilemma $G_3$. In the canonical repeated game, mutual cooperation can be supported for sufficiently large $\delta$ by playing $\{C, D\}$ as punishment profile when player 1 has deviated and $\{D, C\}$ as punishment profile when player 2 has deviated. Then there exists no continuation equilibrium that Pareto-dominates another one. More generally, they show that any $v \in V^*$ is the payoff of a weakly renegotiation-proof equilibrium if $\delta$ is sufficiently large and there exist actions $a^1, a^2$ with $c_i(a^i) < v_i$ and $u_i(a^{-i}) \geq v_i$.

This refinement should – if we can define it for our framework – rule out that both players want to skip the punishment phase of a starting-small strategy. Its adaptation to our framework is somewhat tricky. First, we have to decide when players can negotiate. Second, we have to be specific about what a continuation equilibrium is in our game (note that in our framework, a strategy $\sigma_i$ not only describes how to play in the current group, but also in all future groups).

With respect to the first issue, we rule out negotiations among players who are not matched with each other (there are no information flows between groups). Players can negotiate (1) after being matched together (i.e., after the outcome of the

[9] As Farrell and Maskin (1989) note, the concept of renegotiation-proofness might no longer be appropriate for games with more than two players if they can negotiate in coalitions that are strictly smaller than the number of players.

randomization device is made public), before choosing their action, and (2) after actions have been taken, before choosing between $MT$ and $Q$.

With respect to the second issue, we have to restrict attention to strategies that imply the same path of play in every group. We call a strategy $\sigma_i$ group-independent if (1) the matching decision is always $MT$, and (2) the action choice of a player in role $i$ only depends on the public signals and action profiles realized in her current group.[10] Note that a starting-small strategy is group-independent. We provide a formal definition.

DEFINITION 2 (GROUP-INDEPENDENCE) *A strategy $\sigma_i$ is group-independent if the matching decision is always MT and the action choice in period t only depends on $\{(a^\tau, w^\tau)\}_{\tau\in\{t^*,\dots,t\}}$, where $t^*$ is the last period such that $r^{t^*-1} \neq r^{t^*}$ and $r^\tau = r^{t^*}$ for all $\tau \in \{t^*, \dots, t-1\}$.*

If a strategy $\sigma_i$ is group-independent, it can be described by an *adjunct strategy* $\hat{\sigma}_i$: $\mathrm{H} \times [0, 1] \to \Delta(A_i)$ that maps the history of play in the current group and the current realization of the public signal into the set of mixed actions. This adjunct strategy essentially works like a strategy in the canonical repeated game. It is the same object as in the definition of weak renegotiation-proofness in Farrell and Maskin (1989). In particular, a continuation equilibrium of an adjunct strategy profile $\hat{\sigma} = \{\hat{\sigma}_1, \hat{\sigma}_2\}$ is well defined. Now we can adapt weak renegotiation-proofness to our framework:

DEFINITION 3 (WEAK $gs$-RENEGOTIATION-PROOFNESS) *An equilibrium $\sigma$ in group-independent strategies and its adjunct strategy profile $\hat{\sigma}$ are weakly gs-renegotiation-proof if, whenever players can negotiate, there do not exist continuation equilibria $\hat{\sigma}^1$, $\hat{\sigma}^2$ of $\hat{\sigma}$ such that $\hat{\sigma}^1$ strictly Pareto-dominates $\hat{\sigma}^2$.*

Compared to the definition in Farrell and Maskin (1989), this definition additionally includes the phrase "whenever players can negotiate." The reason for this is that newly matched players can only negotiate in period $t$ after the public signal $w^t$ has been realized. Below we illustrate why this is important.

An equilibrium in starting-small strategies is weakly $gs$-renegotiation-proof only if at the beginning of a relationship at least one player weakly prefers to start play in the punishment phase. The starting-small strategy described at the beginning of this section is therefore not weakly $gs$-renegotiation-proof. Moreover, deviations must be punished in a way so that all players weakly prefer to maintain the relationship. The next example shows that both objectives can be achieved by using the public correlation device.

*Example 3.* Consider again the prisoners' dilemma $G_3$. We construct a starting-small strategy with $\{C, C\}$ as normal-phase profile that supports a weakly $gs$-renegotiation-proof equilibrium when $\delta$ is sufficiently large: each player always

[10] Thus, it does not depend on her period of birth, her identity, the identity of the other group members, or what happened in previous groups (none of these objects exists in the canonical repeated game).

chooses $MT$; if a player starts in a new group in period $t$, profile $\{D, C\}$ is played if $w^t \in [0, 1/2]$, and profile $\{C, D\}$ is played if $w^t \in (1/2, 1]$; after this, profile $\{C, C\}$ is played until a player deviates unilaterally from this profile or the group is broken up exogenously; whenever the player in role 1 (role 2) deviates unilaterally from this path of play (i.e., in the punishment or the normal phase), profile $\{C, D\}$ ($\{D, C\}$) is played in the subsequent period $\tau$ (given that the same group is matched together) if $w^\tau \in [5/8, 1]$, and profile $\{C, C\}$ otherwise; after this punishment, players return to the repeated play of $\{C, C\}$; there is no punishment for bilateral deviations.

We show that this strategy supports a weakly $gs$-renegotiation-proof equilibrium if $\delta$ is sufficiently close to 1. Observe that after any unilateral deviation from this strategy profile, a player either is punished by her opponent or plays in a new group. In the former case, she earns $5/8 \times 4 + 3/8 \times 0 = 5/2$ in expectation in the next period, while in the latter case, she earns $1/2 \times 5 + 1/2 \times 0 = 5/2$. The expected payoffs that occur after the next period are the same in both cases. Hence, she weakly prefers to maintain the relationship after her deviation (though she knows that she will be punished with positive probability in the next period). In each phase, the gain from deviation is 1, while the loss from not playing the normal-phase profile is $4 - 5/2 = 3/2$. Thus, if $\delta$ is sufficiently large, there is no profitable deviation. Note that in each phase and after any signal, at least one player weakly prefers maintaining the current continuation equilibrium to switching to an alternative one. We therefore have derived a weakly $gs$-renegotiation-proof equilibrium in starting-small strategies.

The timing of public randomization is crucial for this solution. It must occur before new groups are matched together. If the players of a new group are able to negotiate continuation play before the outcome of the randomization device is made public, they again could increase expected payoffs by dropping the punishment phase.

We now generalize the example.

PROPOSITION 2 *Assume that there are action profiles $a, a^1, a^2 \in A$ with*

$$c_i(a^i) < u_i(a) \leq u_i(a^{-i}) \tag{1}$$

*and*

$$\max\left\{c_i(a^i) - u_i(a^i), v_i^{\max} - u_i(a)\right\} < u_i(a) - \frac{1}{2}\left(u_i(a^1) + u_i(a^2)\right) \tag{2}$$

*for* $i \in \{1, 2\}$. *For any $\varepsilon > 0$ there is a value $\bar{\delta} < 1$ such that if $\delta > \bar{\delta}$, then a weakly gs-renegotiation-proof sequential equilibrium exists in which the expected normalized payoff for each player in role i is at most $\varepsilon$ away from $u_i(a)$.*

The conditions (1) and (2) guarantee that we can construct a starting-small strategy that has $a$ as normal-phase profile and supports a weakly $gs$-renegotiation-proof equilibrium. The profiles $a^1$ and $a^2$ are played in the punishment phase and after deviations. The condition (1) is the same as in Farrell and Maskin (1989). It requires that we can use $a^i$ to punish a player in role $i$ and at the same time reward her opponent in role $-i$. The condition (2) is new. The left-hand side of (2) is the maximum of what a player in role $i$ gains by deviating from profile $a$, $a^1$, or $a^2$. The right-hand side of (2) is what a player in role $i$ loses when she starts in a new group instead of being in the normal phase, given that $a^1$, $a^2$ are played with equal probability in the punishment phase. The condition (2) is sufficient, but not necessary, for Proposition 2 to hold. Consider, for example, the game $G$ in Figure 6, in which a player in role 1 chooses rows, and a player in role 2 chooses columns. A starting-small strategy with $\{C, D\}$ as punishment-phase profile and $\{C, C\}$ as normal-phase profile can constitute a weakly $gs$-renegotiation-proof equilibrium such that the claim of Proposition 2 holds for $u(a) = u(\{C, C\})$. However, there exist no profiles $a^1$, $a^2$ that satisfy (2) when $u(a) = u(\{C, C\})$.

*Figure 6*
Stage Game $G$

| $G$ | $D$ | $C$ |
|---|---|---|
| $D$ | 0,0 | 2,0 |
| $C$ | 0.9,1 | 1,1 |

In the proof of Proposition 2, we construct a profile in starting-small strategies with the following properties. The punishment phase endures for only one period. The same is true for punishments after unilateral deviations in the punishment phase. Unilateral deviations in the normal phase are punished in a way such that no player strictly wishes to quit the current relationship.

The requirement of weak $gs$-renegotiation-proofness severely restricts the set of action profiles that can be used as normal-phase profile. The maximal gains from deviating in the punishment phase must not be too high relative to the payoff in the normal phase. The following example illustrates this.

*Example 4.* Consider the prisoners' dilemma $G_4$ in Figure 7. A player in role 1 chooses rows, and a player in role 2 chooses columns. It is identical to $G_3$ except that the payoff from mutual defection is 2 instead of 1.

*Figure 7*
Stage Game from Example 4

| $G_4$ | $D$ | $C$ |
|---|---|---|
| $D$ | 2,2 | 0,5 |
| $C$ | 0,5 | 4,4 |

To sustain a weakly $gs$-renegotiation-proof equilibrium in starting-small strategies with $a = \{C, C\}$ as normal-phase profile, only $a^1 = \{C, D\}$ and $a^2 = \{D, C\}$ can be played in the punishment phase. Note that with $a = \{C, C\}$, $a^1 = \{C, D\}$, and

$a^2 = \{D, C\}$ the condition (2) is violated. Assume that $a^1$ and $a^2$ are played with equal probability in the punishment phase and that this phase endures for one period (the argument below extends to any probability distribution and any length of the punishment phase). Let $\delta$ be close to 1, and $M$ large enough that the probability of starting in a new group after choosing $Q$ is close to 1. Now suppose that the action profile $a^1$ is to be played in a new group. By complying with the starting-small strategy, the player in role 1 earns $\approx 0 + 4\delta + 4\delta^2 + 4\delta^3 + \cdots$ in expectation. However, by playing $D$ and $Q$ in the current period and complying thereafter, this player will earn $\approx 2 + 5\delta/2 + 4\delta^2 + 4\delta^3 + \cdots$ in expectation, which is strictly more than the payoff from compliance. Hence, no weakly *gs*-renegotiation-proof equilibrium in starting-small strategies with $a = \{C, C\}$ as normal-phase profile exists for arbitrary $M$.

## 5 Conclusion

We have studied the equilibrium set of infinitely repeated games in which players can quit relationships to find new opponents. We found a structural difference between stage games with two player roles and games with more than two player roles. For two-player games, any individual rational payoff vector in the relative interior of $V^*$ can be an equilibrium payoff for sufficiently large $\delta$. For these games, the option to quit relationships has almost no effect on the set of equilibrium payoffs. By example we showed that such a statement is not possible for games with more than two player roles. Individual rational payoff profiles in the relative interior of $V^*$ may not be equilibrium payoffs. We then translated Farrell and Maskin's (1989) refinement of weak renegotiation-proofness to our framework and provided a characterization of the set of payoffs that can be supported through strategies that are to some extent robust to negotiations between players. In particular, we showed that starting-small strategies that support cooperation in equilibrium can be bilaterally rational in the sense of Ghosh and Ray (1996).

It remains an open question to what extent the option to switch opponents affects the set of equilibrium payoffs. In our framework, players were not allowed to choose their opponents. It may be possible to prove a complete folk theorem when players can force a deviator to remain in the current group and to accept punishment (the model then must specify which option is stronger if one player wants to leave her group while her opponents want the group to stay together). Further research may illuminate which institutions affect the exchange of players between groups and thereby the scope for cooperation.

## Appendix

### A.1 Proof of Lemma 1

Assume that such an equilibrium exists for given $M$ and $\delta$. Then each new player in role $i$ plays $\alpha_i$ at least until she observes that a different profile is played. As

$\alpha \notin \Sigma^{NE}$, there is at least one role $j$ with $c_j(\alpha) > u_j(\alpha)$. Now assume that player $j \cdot m \cdot g$ chooses a best response against $\alpha_{-j}$ in period $t_{j \cdot m \cdot g}$ and plays $\alpha_j$ and $Q$ in each period thereafter. We calculate an upper bound $\bar{N}(\tau)$ on the number of players at the beginning of period $\tau > t_{j \cdot m \cdot g}$ who have observed a profile different from $\alpha$. In period $t_{j \cdot m \cdot g} + 1$ this number is $N - 1$, in period $t_{j \cdot m \cdot g} + 2$ it is $N(N-1)$, in period $t_{j \cdot m \cdot g} + 3$ it is $N^2(N-1)$, and so forth. Thus, we have $\bar{N}(\tau) = N^{t_{j \cdot m \cdot g} - \tau - 1}(N-1)$ for any $\tau > t_{j \cdot m \cdot g}$. Note that in each period, the probability that there are $m \leq M$ players of each role in the pool of unmatched players is at least (when all players choose $MT$)

$$\binom{M}{m}(1-\delta^N)^m(\delta^N)^{M-m},$$

which for given $m$ converges to 0 as $M$ grows large. Thus, the probability that in period $\tau > t_{j \cdot m \cdot g}$ player $j \cdot m \cdot g$ is in a group where profile $\alpha$ is played (given that she has survived up to this period) is at least

$$P(\tau, M) = \sum_{m=\bar{N}(\tau)+1}^{M} \binom{M}{m}(1-\delta^N)^m(\delta^N)^{M-m}\left(\frac{m - \bar{N}(\tau)}{m}\right)^{N-1}.$$

For given $\tau > t_{j \cdot m \cdot g}$ this expression converges to 1 as $M$ grows large. The expected payoff in period $t_{j \cdot m \cdot g}$ for player $j \cdot m \cdot g$ from complying to the equilibrium strategy is

$$u_j(\alpha) + \frac{\delta}{1-\delta}u_j(\alpha), \tag{A1}$$

while the maximal expected payoff from playing a strategy where in period $t_{j \cdot m \cdot g}$ she chooses a best response to $\alpha_{-j}$ is at least

$$c_j(\alpha) + \sum_{\tau=1}^{\infty} \delta^\tau \left(P(\tau, M)u_j(\alpha) + (1 - P(\tau, M))v_i^{\min}\right). \tag{A2}$$

Thus, if $M$ is sufficiently large for given $\delta$, the expression in (A2) exceeds the one in (A1), so that the original assessment cannot be an equilibrium. *Q.E.D.*

### *A.2 Proof of Lemma 2*

Since $v$ is in the relative interior of $V_{NE}^*$, we can find an $\varepsilon > 0$ and a payoff profile $v^l \in V_{NE}$ such that $v_i^l < v_i - \varepsilon$ for all $i$. Then we can choose $T \in \mathbb{N}$ such that

$$v_i - \varepsilon + T(v_i - \varepsilon) > v_i^{\max} + Tv_i^l \tag{A3}$$

for all $i$. Define

$$v^h(\delta) = \frac{1}{\delta^{NT}}v - \frac{1-\delta^{NT}}{\delta^{NT}}v^l. \tag{A4}$$

As $v$ is in the relative interior of $V_{NE}^*$, we have $v^h(\delta) \in V_{NE}^*$ if $\delta$ is sufficiently close to 1. Assume now that there is a pure action profile $a^2$ that generates $v^h(\delta)$ and a Nash profile $\alpha^1$ that generates $v^l$.[11] Let all players play the starting-small strategy mentioned in the text. The expected payoff $E_i$ of a new player in role $i$ is then[12]

$$\begin{aligned}E_i &= \left(1-\delta^{N-1}\right)\left[v_i^l+\delta E_i\right]\\ &\quad+\left(1-\delta^{N-1}\right)\delta^{N-1}\left[v_i^l+\delta v_i^l+\delta^2 E_i\right]\\ &\quad+\cdots\\ &\quad+\left(1-\delta^{N-1}\right)\delta^{(N-1)(T-1)}\left[v_i^l+\cdots+\delta^{T-1}v_i^l+\delta^T E_i\right]\\ &\quad+\left(1-\delta^{N-1}\right)\delta^{(N-1)T}\left[v_i^l+\cdots+\delta^{T-1}v_i^l+\delta^T v_i^h+\delta^{T+1}E_i\right]\\ &\quad+\left(1-\delta^{N-1}\right)\delta^{(N-1)(T+1)}\left[v_i^l+\cdots+\delta^{T-1}v_i^l+\delta^T v_i^h+\delta^{T+1}v_i^h+\delta^{T+2}E_i\right]\\ &\quad+\cdots.\end{aligned}$$

Simplifying this expression yields

$$E_i = v_i^l+\delta^N v_i^l+\cdots+\delta^{N(T-1)}v_i^l+\delta^{NT}v_i^h+\delta^{N(T+1)}v_i^h+\cdots+\frac{\delta-\delta^N}{1-\delta^N}E_i\,.$$

Rearranging this equality gets us

$$E_i=\frac{1-\delta^{NT}}{1-\delta}\,v_i^l+\frac{\delta^{NT}}{1-\delta}\,v_i^h(\delta)\,.$$

The definition in (A4) implies that $(1-\delta)E_i = v_i$ for all $i$. It remains to show that the starting-small strategy supports an equilibrium if $\delta$ is sufficiently close to 1. Whenever profile $\alpha^1$ is played, there is no opportunity to deviate profitably. Assume now that $a^2$ is supposed to be played. The expected payoff of a player in role $i$ from compliance is given by

$$\frac{1}{1-\delta^N}\,v_i^h(\delta)+\frac{\delta-\delta^N}{1-\delta^N}\,E_i\,,$$

while the expected payoff for this player from deviating is at most

$$v_i^{\max}+\delta E_i\,.$$

For $\delta\to 1$ the difference between these two expressions becomes

$$v_i-v_i^{\max}+T\left(v_i-v_i^l\right)>0\,,$$

where the last inequality follows from (A3). Thus, if $\delta$ is sufficiently close to 1, no player can profitably deviate from the starting-small strategy. Finally, all players always (weakly) prefer *MT* to *Q*, since the punishment phase is played at the start of any new group. *Q.E.D.*

[11] If there does not exist a pure action profile $a^2$ that generates $v^h(\delta)$ (a single Nash profile $\alpha^1$ that generates $v^l$), we can replace it with a public randomization with $a^2(w^t)\in A$ ($\alpha^1(w^t)\in\Sigma^{NE}$) for each $w^t$ that yields us an expected payoff of $v^h(\delta)$ (of $v^l$). The proof essentially remains the same.

[12] Without loss of generality we here assume that $T\geq 2$.

*A.3 Proof of Proposition 1*

Since $v$ is in the relative interior of $V^*$, we have

$$u_i\left(\hat{\alpha}_1^2, \hat{\alpha}_2^1\right) \le u_i(\hat{\alpha}^i) < v_i \tag{A5}$$

for $i \in \{1, 2\}$. Choose a small $\varepsilon > 0$ and $T \in \mathbb{N}$ such that

$$v_i^{\min} - \varepsilon + T(v_i - \varepsilon) > v_i^{\max} + Tu_i\left(\hat{\alpha}_1^2, \hat{\alpha}_2^1\right) \tag{A6}$$

for all $i$. We construct a starting-small strategy with $(\hat{\alpha}_1^2, \hat{\alpha}_2^1)$ as punishment-phase profile. If the min–max profiles are not pure action profiles, we can condition the payoffs in the normal phase on the realized action profiles in the punishment phase in a way that makes each player indifferent between her actions in each period of the punishment phase. Define

$$\begin{aligned} v^h(\delta, h^t) = {} & \frac{1}{\delta^{NT}} v - \frac{1-\delta^{NT}}{\delta^{NT}} u_i\left(\hat{\alpha}_1^2, \hat{\alpha}_2^1\right) \\ & - \frac{1-\delta^N}{\delta^{NT}} \sum_{\tau=t-T}^{t-1} \delta^{(\tau-t+T)N}\left[u_i(a^\tau) - u_i\left(\hat{\alpha}_1^2, \hat{\alpha}_2^1\right)\right]. \end{aligned} \tag{A7}$$

Since $v$ is in the relative interior of $V^*$, we have $v^h(\delta, h^t) \in V^*$ for any $h^t \in \mathrm{H}$ if $\delta$ is sufficiently close to 1. Choose a public randomization $a(w^t, h^t)$, which for each $h^t \in \mathrm{H}$ yields an expected period payoff of $v^h(\delta, h^t)$. Consider the following starting-small strategy for players in role $i \in \{1, 2\}$: Always choose $MT$; when you start in a new group, start in phase I below; as long as you are matched to this group, choose the following path of play:

1. Phase I: Choose $\hat{\alpha}_i^{-i}$ in $T$ subsequent periods and then switch to phase II.
2. Phase II: If phase II starts in period $t$, then choose $a_i(w^t, h^t)$ in period $t$ and all subsequent periods as long as there has been no deviation from profile $a(w^t, h^t)$ in period $t$ and ever since. After any deviation, switch to phase I.

The inequality in (A5) ensures that if $\delta$ is sufficiently close to 1, it does not pay for players in role $i$ to play a best response against $\hat{\alpha}_{-i}^i$ in phase I and to quit the relationship (because the highest possible period payoff in the punishment phase is strictly smaller than the expected period payoff in the normal phase, and in a new group phase I would start anew). The equality in (A7) ensures that no player in role $i$ gains by choosing a profile different from $\hat{\alpha}_i^{-i}$ in phase I and to maintain the relationship (because any gain will be wiped out in the normal phase). The inequality in (A6) ensures that there is no profitable deviation in phase II if $\delta$ is sufficiently close to 1. Finally, the equality in (A7) ensures that the expected normalized payoff of a new player in role $i$ is given by $v_i$ if all players comply with the starting-small strategy. *Q.E.D.*

*A.4 Proof of Proposition 2*

Let the action profiles $a$, $a^1$, and $a^2$ with the properties in (1) and (2) be given. For each role $i \in \{1, 2\}$ choose $\bar{w}_i \in (0, 1)$ such that

$$\bar{w}_i u_i(a) + (1 - \bar{w}_i) u_i(a^i) = 0.5 \left(u_i(a^1) + u_i(a^2)\right) . \tag{A8}$$

(1) ensures that this is possible. We now propose a strategy for players in role $i \in \{1, 2\}$ that will support an equilibrium with the desired properties: Choose $MT$ in all periods; when you start in a new group, start in phase I below; as long as you belong to this group, choose the following path of play:

1. Phase I: If $w^t \in [0, 0.5]$, choose $a_i^1$; otherwise choose $a_i^2$. If $w^t \in [0, 0.5]$ and profile $a^1$ has been played or if $w^t \in (0.5, 1]$ and profile $a^2$ has been played, switch to phase II. If the player in role $j$ has deviated unilaterally, switch to phase $\mathrm{II}_j$. If both players have deviated, switch to phase II.
2. Phase $\mathrm{II}_j$: If $w^t \in [0, \bar{w}_j]$, choose $a_i$; otherwise choose $a_i^j$. If $w^t \in [0, \bar{w}_j]$ and profile $a$ has been played or if $w^t \in (\bar{w}_j, 1]$ and $a^j$ has been played, switch to phase II. If the player in role $l$ has deviated unilaterally, switch to phase $\mathrm{II}_l$. If both players have deviated, switch to phase II.
3. Phase II: Choose $a_i$. If profile $a$ has been played or if both players have deviated, remain in phase II. If the player in role $j$ has deviated unilaterally, switch to phase $\mathrm{II}_j$.

Since in each phase a profile from the set $\{a, a^1, a^2\}$ is played, (1) guarantees that in no phase do there exist Pareto-improving joint deviations to another phase. The equality in (A8) ensures that in each phase each player (weakly) prefers $MT$ to $Q$. Thus, if this strategy supports an equilibrium, this equilibrium is weakly $gs$-renegotiation-proof. We now show that this strategy supports an equilibrium if $\delta$ is sufficiently close to 1. The expected payoff $E_i$ of a new player in role $i$ is given by

$$E_i = \frac{1 - \delta^2}{1 - \delta} 0.5 \left(u_i(a^1) + u_i(a^2)\right) + \frac{\delta^2}{1 - \delta} u_i(a) . \tag{A9}$$

In phase I, a player in role $i$ gets, by conforming, an expected payoff of either (case 1)

$$u_i(a^i) + \sum_{\tau=1}^{\infty} \delta^{\tau 2} u_i(a) + \delta(1 - \delta) \sum_{\tau=0}^{\infty} \delta^{\tau 2} E_i ,$$

or (case 2)

$$u_i(a^{-i}) + \sum_{\tau=1}^{\infty} \delta^{\tau 2} u_i(a) + \delta(1 - \delta) \sum_{\tau=0}^{\infty} \delta^{\tau 2} E_i ,$$

depending on the realization of the public correlation device. From deviating in this phase, she gets in case 1 an expected payoff of at most

$$c_i(a^i) + \delta E_i ,$$

while in case 2 she gets an expected payoff of at most

$$c_i(a^{-i}) + \delta E_i .$$

For $\delta \to 1$ the difference between the expected payoffs in case 1 becomes

$$u_i(a^i) - c_i(a^i) + u_i(a) - 0.5\left(u_i(a^1) + u_i(a^2)\right) , \tag{A10}$$

and in case 2 it becomes

$$u_i(a^{-i}) - c_i(a^{-i}) + u_i(a) - 0.5\left(u_i(a^1) + u_i(a^2)\right) . \tag{A11}$$

The assumptions in (1) and (2) guarantee that the expressions in (A10) and (A11) are strictly positive. Thus, if $\delta$ is sufficiently close to 1, no player can profitably deviate in phase I. A similar argument shows that the player in role $i$ cannot deviate profitably in phase $\text{II}_i$ if $\delta$ is sufficiently large. In phase $\text{II}_j$, when $w^t \in (\bar{w}_j, 1]$, the player in role $i \neq j$ gets by conforming

$$u_i(a^j) + \sum_{\tau=1}^{\infty} \delta^{\tau 2} u_i(a) + \delta(1-\delta) \sum_{\tau=0}^{\infty} \delta^{\tau 2} E_i ,$$

while from deviating in this phase, she gets an expected payoff of at most

$$v_i^{\max} + \delta E_i .$$

For $\delta \to 1$ the difference between these expressions becomes

$$u_i(a^j) - v_i^{\max} + u_i(a) - 0.5\left(u_i(a^1) + u_i(a^2)\right) ,$$

which by the assumption in (1) and (2) is strictly positive. Thus, if $\delta$ is sufficiently close to 1, no player can profitably deviate in phase $\text{II}_j$ when $w^t \in (\bar{w}_j, 1]$. In phase II, the player in role $i$ gets by conforming

$$u_i(a) + \sum_{\tau=1}^{\infty} \delta^{\tau 2} u_i(a) + \delta(1-\delta) \sum_{\tau=0}^{\infty} \delta^{\tau 2} E_i ,$$

while from deviating in this phase, she gets an expected payoff of at most

$$v_i^{\max} + \delta E_i .$$

For $\delta \to 1$ the difference between these expressions becomes

$$2u_i(a) - v_i^{\max} - 0.5\left(u_i(a^1) + u_i(a^2)\right) ,$$

which by the assumption in (2) is strictly positive. Thus, if $\delta$ is sufficiently close to 1, no player can profitably deviate in phase II. Finally, phase $\text{II}_j$ is identical to phase II when $w^t \in [0, \bar{w}_j]$. The result then follows directly from (A9). *Q.E.D.*

*References*

Abreu, D., P. K. Dutta, and L. Smith (1994), "The Folk Theorem for Repeated Games: A Neu Condition," *Econometrica*, 62(4), 939–948.

Andreoni, J., and L. Samuelson (2006), "Building Rational Cooperation," *Journal of Economic Theory*, 127(1), 117–154.

Carmichael, H. L., and W. B. MacLeod (1997), "Gift Giving and the Evolution of Cooperation," *International Economic Review*, 38(3), 485–509.

Casas-Arce, P. (2010), "Dismissals and Quits in Repeated Games," *Economic Theory*, 43(1), 67–80.

Datta, S. (1993), "Building Trust," Mimeo, London School of Economics.

Ellison, G. (1994), "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching," *The Review of Economic Studies*, 61(3), 567–588.
Farrell, J., and E. Maskin (1989), "Renegotiation in Repeated Games," *Games and Economic Behavior*, 1(4), 327–360.
Fudenberg, D., and E. Maskin (1986), "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54(3), 533–554.
Fujiwara-Greve, T., and M. Okuno-Fujiwara (2009), "Voluntarily Separable Prisoner's Dilemma," *The Review of Economic Studies*, 76(3), 993–1021.
Ghosh, P., and D. Ray (1996), "Cooperation in Community Interaction without Information Flows," *The Review of Economic Studies*, 63(3), 491–519.
Jackson, M., and A. Watts (2008), "Equilibrium Existence in Bipartite Social Games: A Generalization of Stable Matchings," *Economics Bulletin*, 3(12), 1–8.
Jackson, M. O., and A. Watts (2010), "Social Games: Matching and the Play of Finitely Repeated Games," *Games and Economic Behavior*, 70(1), 170–191.
Kandori, M. (1992), "Social Norms and Community Enforcement," *The Review of Economic Studies*, 59(1), 63–80.
Kranton, R. E. (1996), "The Formation of Cooperative Relationships," *The Journal of Law, Economics, & Organization*, 12(1), 214–233.
Matsushima, H. (1990), "Long-Term Partnership in a Repeated Prisoner's Dilemma with Random Matching," *Economics Letters*, 34(3), 245–248.
Nash, J. (1951), "Non-Cooperative Games," *The Annals of Mathematics*, second series, 54(2), 286–295.
Rob, R., and H. Yang (2010), "Long-Term Relationships as Safeguards," *Economic Theory*, 43(2), 143–166.
Schumacher, H., (2013), "Imitating Cooperation and the Formation of Long-Term Relationships," *Journal of Economic Theory*, 148(1), 409–417.

Heiner Schumacher
Department of Economics and Business
Aarhus University
Fuglesangs Allé 4
8210 Aarhus V
Denmark
hschumacher@econ.au.dk